Verb Errors in Advanced Spoken English

Article *in* AUC PHILOLOGICA · July 2017

DOI: 10.14712/24646830.2017.8

CITATIONS
2

READS
2,325

1 author:

Tomáš Gráf
Charles University in Prague
42 PUBLICATIONS **33** CITATIONS

SEE PROFILE

2017 ACTA UNIVERSITATIS CAROLINAE PAG. 131–149
PHILOLOGICA 1 / PRAGUE STUDIES IN ENGLISH

VERB ERRORS IN ADVANCED SPOKEN ENGLISH

TOMÁŠ GRÁF

ABSTRACT

As an experienced teacher of advanced learners of English I am deeply aware of recurrent problems which these learners experience as regards grammatical accuracy. In this paper, I focus on researching inaccuracies in the use of verbal categories. I draw the data from a spoken learner corpus LINDSEI_CZ and analyze the performance of 50 advanced (C1–C2) learners of English whose mother tongue is Czech. The main method used is Computer-aided Error Analysis within the larger framework of Learner Corpus Research. The results reveal that the key area of difficulty is the use of tenses and tense agreements, and especially the use of the present perfect. Other error-prone aspects are also described. The study also identifies a number of triggers which may lie at the root of the problems. The identification of these triggers reveals deficiencies in the teaching of grammar, mainly too much focus on decontextualized practice, use of potentially confusing rules, and the lack of attempt to deal with broader notions such as continuity and perfectiveness. Whilst the study is useful for the teachers of advanced learners, its pedagogical implications stretch to lower levels of proficiency as well.

Keywords: advanced spoken English, learner corpora, error analysis, tenses, verbs, present perfect

1. Introduction

The Common European Framework of Reference (Council of Europe, 2001, pp. 24–27) defines the advanced levels (i.e. C1 and C2) of language proficiency in its Global Scale and self-assessment grids using such terms as coherence, spontaneity, fluency, precision, complexity, flexibility, efficiency, and effortlessness. It further describes advanced accuracy as being characterized by a high degree of grammatical accuracy, consistent grammatical control and rare occurrence of errors. It is the experience of many a language teacher that whilst the speech of truly advanced learners displays all of the above mentioned characteristics and is certainly not riddled with errors, occasional instances of inaccuracies still occur, and sometimes it might even surprise the hearer how seemingly basic these

DOI: 10.14712/24646830.2016.24

appear to be. As has been shown, for example by Granger (1999), Götz (2015), and Gráf (2015a), such errors often involve the use of tenses. It is the aim of this study to explore the nature of grammatical errors in the use of verbs in spontaneous spoken production of Czech advanced learners of English. At the time of writing no recent analysis of such material is available.

1.1 Error analysis and its offshoots

It might seem somewhat anachronistic to revert our attention to errors decades after the gradual decline of Error Analysis and at a time when the increased use of English as a lingua franca and the communicative language teaching methodologies call for increased tolerance to less than perfect L2 performance. The recent advances in the field of Learner Corpus Research (see especially Granger et al., 2015) and the development of Contrastive Interlanguage Analysis (Granger, 1996) and Computer-aided Error Analysis (Dagneaux et al., 1998) have made available not only new sources of data but also new analytical and interpretative techniques. Together these can yield deeper insight into the nature of L2 production and acquisition and can inform the fields of language acquisition and pedagogy. Last but not least new research in this area within the framework of English as a foreign language (EFL) may show to what extent the original findings are applicable to the English spoken by a generation of young learners (Meriläinen, 2010) who have had entirely different opportunities throughout the learning process than the preceding generations, e.g. early starting age at schools, opportunities to practise and travel, easy access to ELT and authentic materials, significant exposure stemming from the use of English as international language on the internet and elsewhere, and this list could easily continue. Aided by the status of English and its omnipresence, large numbers of learners nowadays frequently attain much higher levels of foreign language competence than is customary in other foreign languages than English. This gives researchers the opportunity to study advanced language proficiency and answer questions regarding fossilization and maximum L2 attainment. It is in this area that the study of errors is especially useful even though many of the methodological problems identified in the early days of contrastive and error analysis – such as the very definition of error – remain the same.

The systematic study of language-learner errors became the preoccupation of applied linguists in the late 1950s when Lado (1957) laid the foundations for the so-called Contrastive Analysis Hypothesis which claimed that the main difficulties in acquiring a foreign language were caused by differences between the learner's L1 and the target language. Such difficulties were considered predictable if a methodical contrastive analysis of the two languages in question was carried out, and consequently such pedagogical procedures could be designed as would target the problematic aspects in order to prevent language errors. The clearly aberrant belief that language teaching was primarily a question of error prevention and the gradual realization that learners did not behave exactly in the way contrastive analysts had predicted inevitably led to the decline of the methodology. The interest in errors, however, persisted. A new methodology, Error Analysis (EA) (Corder, 1967, 1974), studied errors not as obstacles in language learning but as evidence of learners' internal hypotheses regarding the target language system. Such

information would prove invaluable for the understanding of L2 acquisition and production. EA provided typologies of errors and linked them to the different L2 developmental stages whilst moving away from the questionable attempts to identify sources of errors. One of EA's most important contributions was the differentiation between systematic and performance errors, which proved useful to both language teachers and learners.

However, there were several methodological problems. The key one was the inability to provide a fool-proof definition of error. Definitions mostly juxtaposed errors with norms by claiming that errors were deviations from the norm, or even more loosely that errors are "unsuccessful bits of language" (James, 1998, p. 1). The problem instantly arose as to which norm to work with and in which different contexts. What might be incorrect in one particular situation could be entirely "normal" in another, and the whole picture might receive different contours in different varieties of the target language, a realization which led Dušková (1969) to point out that there was a continuum between deviation and acceptability, or, as Gilquin and De Cock (2011) observed, between error and dysfluency.

The waning interest in errors was rekindled in the 1990s with the development of computerised learner corpora, which resulted in many pedagogical applications such as usage notes and error warnings in learner dictionaries (e.g. Rundell, 2007), a dictionary of errors (Turton and Heaton, 1996) and Swan's and Smith's (2001) treatise of learner English around the world. The study of errors also prepares the ground for the development of accuracy studies within the CAF (complexity, accuracy and fluency) model of language proficiency and production (Housen et al., 2012).

1.2 Verbal categories in learner English

Verbal categories include number, tense and aspect, mood, and voice. Of these it is tense and aspect that are of particular interest to language teaching. Both tense and aspect are verbal categories of semantic nature. Tense is a type of deixis whose function is to provide temporal reference, or – in other words – relate the actual time of the event, action or state described by the verb to some other time, most commonly the time of speaking. Aspect expresses the stance the speaker takes to the progress of the described event and encodes among other such notions as completeness, continuity, sequentiality, intentionality, and iterativeness.

The use of tenses and aspect is a notoriously difficult area for the learners of foreign languages. English, despite its proverbial simplicity, is not an exception. This is especially true for learners whose own mother tongue's tense and aspect system is as different as in the case of Czech learners learning English. For the ease of teaching, pedagogical grammars of English traditionally subsume the category of aspect under the broader category of tense and differentiate between present, past, future, present perfect, past perfect and future perfect, and within each of these further distinguish between simple and continuous/progressive forms. Thus English textbooks and pedagogical grammars recognize 12 different tenses. Once they are labelled in this way, they are usually introduced separately and then comparisons are made between simple and continuous forms of each tense. This practice does not encourage the development of the awareness of aspect; teachers teach present simple, then present continuous, but they do not deal with the concept of continuity on its own. Similarly, they do not deal with the concept of

perfectiveness but teach the perfect tenses as separate entities not related to each other. Moreover, they frequently take sentences out of context and rather than develop in their students the feel for the broader notions and the roles they play in the construction of discourse, they "equip" the learners with often confusing rules such as "in sentences with *since* always use the present perfect". Consequently, many learners come to believe that the choice of a grammatical aspect depends on the presence of an overt feature (such as a particular adverb).

The basis for the distinction between the tenses derives from morphological markers but it does not deal with broader notions. Thus the present perfect is identified as a form consisting of an auxiliary *have* and the past participle of the lexical verb, and then rules are given for the different uses of the tense. The very system of the twelve tenses thus inherently encourages deductive and present-practice-produce approaches to the teaching of grammar at the cost of induction and discovery which promote the development of linguistic awareness and from a pedagogical point of view appear more logical.

The teaching of tense and aspect in English to beginners usually commences with the present simple forms of the verbs *be* and *have*. These are usually acquired rapidly with only occasional mistakes in using the form *have* for *has* in the 3rd person singular. Lexical verbs are gradually introduced to be used in the present simple. The acquisition of the third person *–s* ending is, however, slow and problematic despite its conceptual simplicity, and slips when the student fails to produce the form may occur even beyond the intermediate level. Interestingly, its acquisition has been shown to be slow even for children acquiring English as L1 when the plural *-(e)s* ending and the genitive *'s* tend to be acquired first (Clark, 2009; Ingram, 1989). This is probably caused by the relative infrequency of the ending (compared to the higher frequency of the plural marker, and the frequency with which we refer to personal possessions when communicating with little children, e.g. *Mummy's bag, Daddy's car* etc.) and by its superfluity from the communicative point of view.

The present continuous is introduced soon and it has been pointed out (Conrad, 2016) that its teaching and practice takes precedence over the present simple even though it is the latter that appears in English with greater frequency. The present continuous is generally taught as a structure with concrete rules for use whilst the concept of continuity and the reasons for making it explicit are usually ignored. This may cause problems for speakers of languages (such as Czech and German) which do not make this distinction.

As much as continuity as a broader concept is ignored, so is perfectiveness. Explanations tend to focus on individual types of usage and many "helpful" rules are introduced when the students are advised to choose the tense if the sentence contains a particular, overt feature. This approach is then further strengthened by decontextualized practice which makes use of these overt features, rather than by contextualized, discourse-based approaches.

It is of no surprise that studies of learner language find the areas of tense and aspect the most error-prone (Davydova, 2011; Dietrich, Klein, and Noyau, 1995; Eriksson, 2004, 2008; Götz, 2015; Gráf, 2015a; Granger, 1999; Hinkel, 2004; Rogatcheva, 2009, 2012; Salaberry and Shirai, 2002).

Other morphological categories of verbs – number, mood, and voice – with regard to spoken learner language and related errors have been studied less extensively. Numer-

ous studies have dealt with the use of active and passive voice in writing (e.g. Granger, 2013). Likewise, modality has drawn some attention (e.g. Guo, 2005) and learners' use of the conditional has been explored for example by Götz (2015). Whilst the current study explores errors in these categories as well, its focus is on tense and aspect; as is about to be shown, they vastly outnumber the other instances of grammatical inaccuracies.

2. Material and method

The material for our study draws on the Czech subcorpus of LINDSEI[1] (Gilquin et al., 2010). It will be further referred to as LINDSEI_CZ (Gráf, 2015b). LINDSEI is a multi-national corpus of advanced spoken learner English. At the time of writing it comprises 20 national subcorpora, each with at least 50 speakers' recordings with average duration of 15 minutes and corresponding orthographic transcriptions. LINDSEI was conceived as the spoken counterpart of the written corpus ICLE (Granger, 2009).[2] The speakers perform three tasks (a monologue on a preselected topic, a dialogue and a picture-based narrative). The speakers' proficiency had not been established prior to their recruitment. Instead, a method of institutional selection was adopted and LINDSEI recruited speakers from among students of English philology in at least the third year of their university course. It has been shown (Götz, 2015; Gráf, 2015) that this is a weak point in the corpus design as proficiency within the corpus can vary (Carlsen, 2012). Some of the national subcorpora (French, Taiwanese, Czech) have consequently engaged professional proficiency raters, but to date only the French subcorpus has been fully assessed, the work on the other two being in progress.

For this study spontaneous speech production has been chosen as it is here, under the pressure of online planning, on the assumption that the learners are likely to slip more often than in writing, where a much higher degree of accuracy might be expected as a result of the extended time for planning.

Table 1 provides an overview of LINDSEI_CZ as regards the number of speakers, the length of the interviews, the number of the tokens, and basic metadata.

Table 1. LINDSEI_CZ – the volume of recorded and transcribed data

Choice of topic for Task 1	**Length of A & B turns[3] in tokens**	**Length of B turns only in tokens**	**Duration of A & B turns (hh:mm:ss)**	**Duration of B turns (hh:mm:ss)**	**Mean length of interview in tokens**	**Mean duration of interview (mm:ss)**
Country = 22 Film/play = 18 Experience = 10	123,761	95,904 mean = 1,918 (SD = 407)	12:52:25	10:37:42	2,475 (SD = 386)	15:27 (SD = 2:14)

[1] The Louvain International Database of Spoken English Interlanguage
[2] The International Corpus of Learner English
[3] The phrase "A turn" and "B turn" denote utterances made by the interviewer and the interviewee, respectively.

The method used is Computer-aided Error Analysis (Dagneaux et al., 1998). As a first step, errors were identified[4] in the transcribed corpus and annotated using a system of descriptive, incremental tags described in the Louvain Error-tagging Manual (Dagneaux et al., 2008). A typical tag has a number of positions which are occupied by letters denoting the particular type of error. The first position is the most general (for example the letters L and G denote lexical and grammatical errors), the subsequent positions add further detail (e.g. V in the second position stands for verb, and in the third position for voice). Thus GVT stands for **G**rammar-**V**erb-**T**ense and GVV for **G**rammar-**V**erb-**V**oice errors. The system can be easily adapted to suit the design of the study and the researcher's intentions. Thus Kämmerer (2009) extended the original tag LS used for marking Lexical Single errors to LSP to mark Lexical Single Prepositional errors. The tags are entered in round brackets.

As a second step, the target hypothesis is entered after the erroneous form. The target hypothesis is a suggestion made by the error annotator as to what he assumes the speaker was intending to produce. In other words, it is what the annotator assumes to be the correct or more appropriate form. This serves simply as a guide to the annotator's way of thinking, and hence only one target hypothesis is suggested although different options might be possible. Target hypotheses are entered between dollar signs. The following examples present two utterances containing tagged errors and target hypotheses. Example (a) shows a grammatical error in the use of a verb tense, example (b) then a lexical phrase (LP) error. In all of the examples in this study the examples are directly lifted from the transcribed corpus LINDSEI_CZ including the various transcription conventions (such as the dot denoting a silent pause). No punctuation is used in the transcriptions.

(a) the whole time it **(GVT) hasn't rained $didn't rain$** one day
(b) I think that I would **(LP) say . truth $tell the truth$**

As a third step all GV (Grammar-Verb) errors were extracted using simple queries in the concordancer AntConc (Anthony, 2014). The examples were then sorted, further coded and analysed using the data functions of MS Excel.

3. Results

The error identification and tagging process revealed a total number of 250 errors involving morphological aspects of verb usage. This constitutes 19.2% of the total number of errors (n=1,301) identified in LINDSEI_CZ (see Gráf, 2015b). In the whole corpus only 4 speakers did not commit any of these errors.

As expected, errors in the use of tenses (including tense agreement) are most frequent, forming 69.2% of all verb errors. These errors were committed by 42 (84%) speakers. The second group of most frequent errors is made up of errors in the use of the conditional,

[4] Problematic cases were compared with similar utterances in the parallel corpus of spontaneous native speech LOCNEC. The comparison revealed that, for example, the use of the past perfect is common in native speech even in situations where its use is not entirely necessary as the sequence of events is clear from the context.

with 24 errors (9.6%) and 14 (28%) erring speakers. The remaining types are considerably less frequent (see Table 2).

Table 2. Verb error frequencies, numbers of erring speakers and mean numbers of errors

Type	Count	%	Number of erring speakers (n=50)	Mean number of errors per speaker	SD
GVT – verb tense errors (including tense agreement errors)	173	**69.20%**	42 (84%)	**3.46**	3.21
GV*C – errors in the use of the conditional	24	**9.60%**	14 (28%)	**0.48**	1.04
GVM – erroneous forms	20	**8.00%**	11 (22%)	**0.4**	1.22
GVAUX – erroneous use of auxiliary or modal verb	17	**6.80%**	14 (28%)	**0.34**	0.68
GVN – verb number errors	14	**5.60%**	10 (20%)	**0.28**	0.60
GVV – erroneous use of voice	2	**0.80%**	2 (4%)	**0.04**	0.20
Total	250	**100.00%**	**46 (92%)**	**5**	4.05

The mean numbers of errors and the high standard deviations (SD) show that the performance of the individual speakers varies significantly: for some of the categories only a small proportion of the speakers were involved and in most categories there appeared speakers who committed more errors in the individual categories than the others. This is more clearly visible from Figure 1 which shows the relevant boxplots.

Table 3 shows the frequency of errors in the use of the individual tenses. As was expected, the most problematic tense is the present perfect; it was erroneously used in 37 cases (28.46% of all GVT errors), and not used when it was supposed to in 58 cases (44.61%). Altogether then the present perfect (wrongly used or wrongly not used) appears in 73% of all verb tense errors. The following largest group is formed by errors in the use or non-use of the past simple, which is incorrectly used in 69 cases, and incorrectly not used in 37 cases. There also appear 20 cases of erroneous use and 8 cases of non-use of the present simple. The past perfect was erroneously used in 1 case, and not used when it was preferable in 24 cases. There are 3 cases of errors in the use of the present continuous, and 1 instance of wrongly used future simple. Errors in other tenses such as the future continuous or perfect have not been identified.

3.1 Present perfect errors

As has been shown above, the present perfect appears to be the most problematic tense for the speakers in LINDSEI_CZ. The 37 examples of its incorrect use were produced by 20 (40% speakers). The 58 instances in which a different tense was used erroneously instead of the present perfect were produced by 28 (56%) speakers. Altogether 34 speakers (68%) either produced at least one instance of incorrect use of the present perfect or at least one instance when they erroneously used a different tense instead. The following section provides an overview and typology of these errors.

Figure 1. Boxplots showing the distribution of errors in the following categories: verb tense including tense agreement (GVT_A), conditional (GV_C), form (GVM), voice (GVV), number (GVN), auxiliary and modal verbs (GVAUX). The last boxplot (Total) provides a summary of all of the instances.

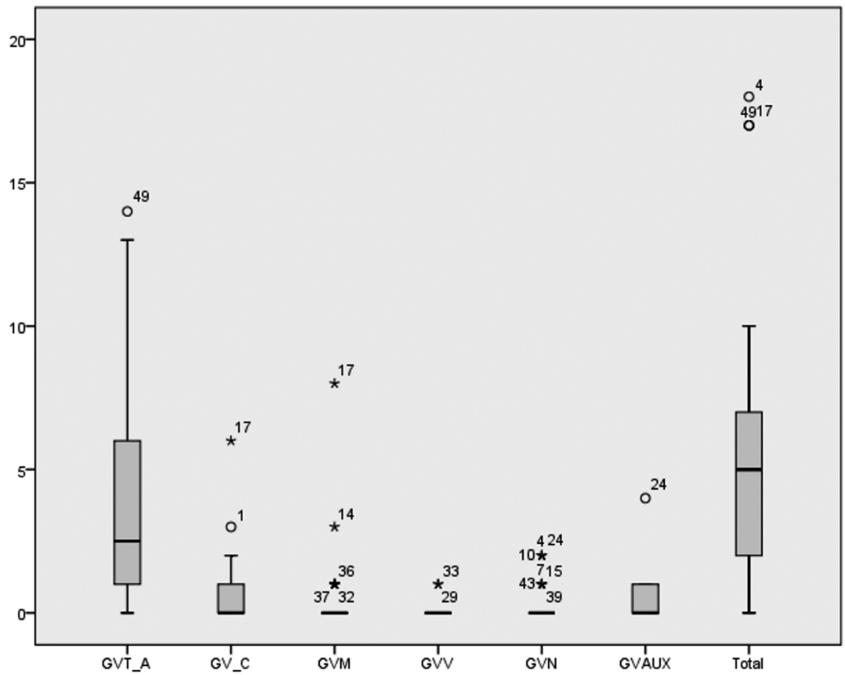

Table 3. Pivot table listing the frequency of erroneously used tenses (rows) and the hypothesised tenses (in columns)

Erroneously used/ Corrected	Past continuous	Past perfect	Past simple	Present continuous	Present perfect	Present perfect continuous	Present simple	Total
Future simple							1	**1**
Past perfect					1			**1**
Past simple	2	13			48	1	5	**69**
Present continuous					1		1	**2**
Present perfect		10	25				1	**36**
Present perfect continuous			1					**1**
Present simple		1	11	1	6	1		**20**
Total	**2**	**24**	**37**	**1**	**56**	**2**	**8**	**130**

3.1.1 Erroneous use of the present perfect

Let us first explore the instances in which the present perfect was used in place of a more appropriate tense (25 instances, 16 speakers). The tense was used despite the fact that specific time or occasion were explicitly mentioned as part of a larger narrative scheme thus requiring the use of the past tense or past perfect. A closer analysis of the examples reveals that there might exist a number of features which might trigger off the production of the present perfect even in situations which have no relation to the present. Such triggers are either typical situations in which the present perfect is used when it does refer to the present, or other lexical or grammatical features which the learner associates with the present perfect.

Such is the situation in which it is used to refer to recently completed actions in which the speaker expresses a sense of achievement (there were 4 such cases in our corpus). This is illustrated by examples (1) and (2) which were both part of a longer past-tense narrative.

(1) well first of all (GVT) I've learned $I learned$ (er) . to speak English very well there
(2) I think (GVT) it's encouraged $it encouraged$ me to to (er) do that

A similar situation (there were 3 cases in our corpus) may be observed in examples (3) and (4) in which the trigger is the sense of experience (which is even explicitly mentioned in example 4).

(3) after after high school […] (GVT) I've made $I made$ a wrong decision
(4) when I was seventeen . and (eh) (GVT) it's been $it was$ (em) great experience

Example (5) illustrates a type (there were 3 cases in our corpus) where the speaker expresses duration, which is often a concept learners associate and frequently practise in the present perfect.

(5) because (GVT) I've been $I was$ there for a year which was (eh) . a long time

Examples (6–8) illustrate how verbs commonly used in the present perfect in sentences referring to experience may act as its trigger even in sentences with past reference without any relation to the present. It may be hypothesised that such verbs are frequently used in sentences for present perfect practice. Our sample contains 7 such instances of the verb *see,* 2 cases of *visit*, and one of *meet.*

(6) (GVT) I've seen $I saw$ it (er) on my birthday
(7) (GVT) I've visited $visited$ (eh) Portugal . Lisbon (eh) this summer
(8) and secondly . (GVT) I've met $I met$. (er) many interesting people

Example (9) contains the intensifier *just* which is again frequently used in the present perfect and in practice sentences.

(9) they were incredibly (er) forthcoming . to us and (er) (GVT) it's just been $it was$ just wonderful

Example (10) appears to be another typical textbook and practice context for present perfect practice.

(10) in my bachelor studies (GVT) I've been studying $I studied$ also (er) . also literature

In this group there are also cases, such as example (11), where no likely trigger can be identified.

(11) and then (GVT) I've thought $I thought$ it would be great to to to be able to engage with

3.1.2 Failure to supply the present perfect

A large group of errors (58 instances, 28 speakers) revealing the complexity of the tense to the learners is made up of instances where the learners failed to produce the present perfect when it was due, replacing it by a different tense. Whilst in the previous section we could hypothesise about possible triggers (mainly teaching- or analogy-induced), the examples in this section are symptomatic of insufficient understanding or automatization of the rules, and some of the errors might be the result of L1 transfer. This is common for example in sentences which express duration at the present moment (see examples 12 and 13) which Czech expresses using the present simple, and also very commonly in complex clauses containing the conjunction *since* (see examples 14 and 15).

(12) it is because I (GVT) teach $have been teaching$ the girl for about . three years already
(13) but it it (GVT) it's $it's been$ open for just a few years
(14) since that moment I . I (eh) (GVT) wished $have been wishing$ I could return .
(15) ever since I was . thirteen or fourteen I I (GVT) had $have had$ a . list . of countries I w= I . want to . visit

Further errors in this group include those where the present perfect was not used for the result of a recent activity (ex. 16). Example (17) presents a common usage in American English and if it were not for the speaker's entirely British-like accent it would have to be excluded from our account.

(16) that's how he . (em) how he (GVT) drew $has drawn$ her and how he probably sees her
(17) like she (GVT) just came $has just come$ from her hairdresser's

Unlike in examples (6–8) which illustrated a common situation in which the speaker uses the present perfect to refer to experience which is located in a concrete situation or context, examples (18–21) show the opposite: here the students fail to use the present perfect when referring to a past experience whose effect is currently relevant.

(18) and I (GVT) was (er) in $have been to$ (eh) Germany twice
(19) it's written by a German author but . I (GVT) didn't read $haven't read$ it
(20) yes I (GVT) did $have done$ it before but you know this is like a sort of different experience because
(21) I've tried some translation I've I've I (GVT) did $have done$ some little little teaching

Considering the high proficiency of the speakers, these errors are somewhat surprising and clearly show that the automatization of the control of a complex tense which has no one-to-one parallel in the speakers' mother tongues is a slow process. It is, however, also possible that these instances are mere slips and we would have to take into account not only the erroneous instances in our corpus but also all of the correct ones. Similarly, we would not expect problems in sentences referring to an event started in the past and not finished at a present point in time, as this usage is usually taught early on in language courses and textbooks and revised frequently. Examples (22–23) show that errors do occur here as well, and we can only speculate if they are the result of L1 transfer, as Czech would use the past simple in the same situations.

(22) well I (GVT) always wanted to $have always wanted to$. (erm) .. so I hope . I can become a teacher but
(23) I don't think she (GVT) ever s= . stopped $has ever stopped$. as an au pair . I think she still . (eh) looks after

The last group of examples in which students incorrectly used a different tense other than the present perfect has altogether 31 examples produced by 18 speakers. They are sentences in which the present perfect refers to a past or recent event which has direct consequences at present. Their high incidence suggests (and the many years of my teaching experience support this view) that this usage of the present perfect poses great problems for the learners as they have to make a choice of the tense not only based upon the event's temporal reference but also on the judgement of the current relevance. The difficulty might arise from the fact that Czech does not make this distinction and simply uses the past tense for such examples.

(24) and he comes back .. and . all his arts collection everything is gone it (GVT) was $has been$ stolen
(25) in the past year (erm) . I I (GVT) changed $have changed$ a lot
(26) the picture is finished so he (GVT) showed $has showed$ her the the final version and she . seems . looks quite surprised
(27) I don't know why I just (GVT) didn't develop $haven't developed$ any relationship with it

3.2 Errors in the use of other tenses and aspect

Besides the group of 95 errors in the use of the present perfect, our corpus contains 35 examples of sentences containing erroneous use of other tenses produced by 20 speak-

ers. The largest group is made up of instances of erroneous use or non-use of the past perfect (altogether 23 errors committed by 14 speakers). Most frequently the past simple is used instead, which makes the actual sequence of events rather unclear. In example (28) the speaker is referring to problems she had experienced prior to the trip which she is describing, but the lack of the past perfect might imply the problems had been experienced during the same trip. Similarly, the use of the past perfect would have prevented the ambiguity as to the temporal interpretation present in example (29), where we are left wondering whether the portrayed person was unhappy with the result or the process of being painted.

(28) there had to be plenty of water because we (GVT) experienced $had experienced$ some problems with water
(29) so probably she's not happy with the way she (GVT) was $had been$ painted

There appear to exist a number of triggers leading to the failure to produce the past perfect. This is true of the cases in which it is replaced by the present perfect,[5] as in examples (30–33). Such triggers include the chunk *been to*, the verbs *read* and *see* and the adverb *just*, all of which are frequently used during present perfect practice.

(30) I tried to travel as much as I could because (eh) (GVT) I've never been $I had never been$ to Ireland before
(31) and it it was amazing I (GVT) I've read $I'd read$ the play before though so that might have made it
(32) it was my dream to see it because I (GVT) have seen $had seen$ many films and and serials about it
(33) I'd never been to Egypt . and I: . (GVT) I've $I'd$ just (erm) . been reading many books about it

Examples of errors in the use of other tenses are much less frequent and involve fewer speakers. Two speakers use the future simple in a conditional clause after the conjunction *if,* as in example (34). This is possibly caused by L1 transfer (cf. the Czech *jestli budu přijat*). The remaining examples are most likely slips which are difficult to categorize and explain, as in examples (35–36).

(34) hopefully I'd like to if if (GVT) I'll be $I am$ admitted
(35) I already (GVT) know $knew$ them very well
(36) but . I wasn't that interested I just . I (GVT) like $liked$ the social studies

Only four speakers erred in the selection of aspect, using continuous for simple, as in example (37), or simple where continuous would appear more natural, as in examples (38) and (39), which are actually harder to interpret as a result.

(37) while she oh she's actually (erm) having a squint $she actually has a squint$

[5] The frequency of these examples is included in section 3.1.1.

(38) she didn't know (erm) what she (GVTA) signed $was signing$
(39) and they they told us when we (GVT) left $were leaving$ you know you should come to visit us . much more often

3.3 Erroneous sequence of tenses in reported speech and dependent clauses

In our corpus there are also 43 errors (committed by 22 speakers) in the use of tenses in dependent clauses following a main clause in the past tense. From a pedagogical point of view, this is another area of great difficulty for Czech speakers of English and one frequently practised and taught as the tense in Czech dependent clause is not affected by the tense in the main clause. In our system of classification, these errors form a subgroup of the grammar-verb-tense errors and are tagged as GVTA, where the A stands for *agreement,* which is a label commonly used in language textbooks (i.e. tense agreement).[6] Eleven examples occur in reported speech after verba dicendi. These include the verbs *ask, say,* and *tell* and also the verb *text* (as in *to write in a text message*) (see examples 40–42).

(40) when I came back my mum (er) told me that (er) she (GVTA) expected $had expected$ something like that to happen
(41) when I was standing here my father called me and asked me how I (GVTA) am was
(42) I just texted her (GVTA) I'm $I was$ coming home three days (er) (er) earlier

Other verbs include *think* (6 times), *feel* (5 times), *realize* (3 times) and single instances of *agree, decide, figure out, find out, check, know, make clear, see, suppose, understand, want to know/show*, and the copular *be (surprised)* and *be (shocked)* (see examples 43–46). Example (46) illustrates that such an error can actually affect the meaning and interpretation of the whole sentence; the one produced by our speaker would imply that the decision in question had not yet been made, whilst the opposite (as suggested by the target hypothesis) is actually the case.

(43) I thought that I (GVTA) can $could$. see the sea (eh) or ocean (eh) maybe it was just my impression
(44) I was there for four months and so in the middle of it I really felt I (GVTA) want $wanted$ to leave
(45) he made it quite clear that he (GVTA) doesn't $didn't$ really like the other boy
(46) the other women agreed that this (GVTA) was $had been$ entirely her decision

Further 2 errors have been recorded in the sequence of tenses in clauses of comparison, as exemplified in (47).

(47) when he would . paint her as he . (eh) (GVT) sees saw sees her

[6] Another commonly used label is sequence of tenses, or in Czech časová souslednost.

3.4 Conditional errors

Besides tense errors, the recorded speakers also committed errors in the use of the conditional. There were a total of 24 of these errors, and they were committed by 14 speakers. They include mainly the wrong choice of mood in conditional clauses introduced by *if*, as in example (48), and instances of failing to use the perfect infinitive form in the if-clause when referring to a past event, as in example (49). Whilst the conditional is frequently practised in language classes, experience tells me that especially the latter error is fairly frequent even with advanced learners. This might be the result of L1 transfer, as contemporary informal Czech tends not to use the past conditional. This also applies to the frequent error in the failure to use the past perfect when referring to a hypothetical past event, as in example (50).

(48) if . (er) the woman . (er) (GVAUXC) would be $were$. (er) a really good friend
(49) and it would (er) (GVAUXC) cause $have caused$ me like (er)
(50) if they (GVTC) stayed $had stayed$ over there and in the place where they were so they would have probably become

3.5 Other types of verb errors (auxiliaries, number, voice, form)

For convenience sake, the remaining 4 groups of errors are dealt with in one section, including the group of 17 errors in the use of auxiliary and modal verbs (tagged GVAUX), examples 51–54), 14 errors in the verb number category (tagged GVN, examples 55–57), 2 examples of errors in the use of grammatical voice (tagged GVV, example 58).

(51) we . weren't in a hotel .. so . I (GVAUX) 0 really looking $was really looking$ forward
(52) all the weeks of the study . (eh) including holidays . (er) we (GVAUX) should have read $were supposed to read$
(53) I always try to google it so that I (GVAUX) would 0 remember in the future
(54) there is nothing to carry sound and so you . of course you (GVAUX) 0 can hear the voices of the characters
(55) everyone in (erm) (eh) . in my surrounding . around me . (GVN) know $knows$ English
(56) maybe the hair . the the her hair (GVN) are is different
(57) an experience that (GVN) have has . taught me something
(58) when the children are running to (eh) (GVV) get hidden $hide$ somewhere in the garden

The examples in these groups are so few that it is impossible to work out any patterns or areas of particular difficulty: in the GVN group there appear 9 instances of the wrong form of the third person singular verb (missing *-s*). Example (52) shows a fairly common error in translating the Czech *měli jsme* as *we should have done* rather than *we were supposed to do.* Transfer is most likely at play in example (53) where in the dependent

clause Czech uses conditional while English uses indicative, and in example (56) where the erroneous use of the plural form of the verb was most likely the result of the Czech equivalent for *hair* being the plural noun *vlasy*.

The last group of errors is difficult to categorize; they include odd instances of slips (tagged GVM) in the use of the correct form (example 59) or even in the use of negation (ex. 60). The low frequency of these errors shows that they are not systematic and do not require pedagogical intervention.

(59) case . (er) it was very interesting for me to (GVM) found $find$ out that actually
(60) yeah I am but I (GVM) no study $don't study$ (LP) English language

4. Discussion

One of the key concerns of this study was the exploration of spontaneous speech produced by advanced learners with the view to ascertaining the key types and frequency of verbal errors and thus identifying those grammatical aspects that would appear most problematic. The results have shown that despite their high proficiency nearly all of our advanced learners make occasional or even more regular slips in the area of verb usage, and this is especially true of the use of tenses. But even here there appear areas of particular difficulty, which is especially the use of the present perfect, and tense agreement.

The results are similar to Granger (1999), and especially to Götz's (2015) study of grammatical errors produced by the participants in the German subcorpus of LINDSEI. LINDSEI_CZ and LINDSEI_GE are very similar in size, the total number of errors identified (1,301 and 1,335 errors respectively) and their distribution. Götz identified 235 GVT errors, which make up 17.6% of the total number of errors in LINDSEI_GE. The proportion of GVT errors in the LINDSEI_CZ is 13.3%.

The major difference between the two corpora is the larger number of aspect (confusing simple and continuous) errors in the German subcorpus (30.3% of all GVT errors in German as opposed to 3.2% in Czech), and the disparity between the frequency of errors involving the present perfect (38% of all erroneous instances in Czech as opposed to 28% in German).

Similarly to our finding, Götz recognizes negative transfer from the L1 as one of the key factors causing the production of errors. A deeper comparison of these two similar studies cannot be performed at present, as Götz's classification of errors worked with a somewhat courser typology of errors (e.g. not dealing systematically with tense agreement errors).

Likewise, Davydova (2011) in her complex study of the present perfect produced by speakers of a variety of L1s finds similar causes for the difficulty these speakers experience – mainly the absence of the tense in the L1 (as in the case of Russian speakers) and the consequent attempt to express the notion in question using structures the speaker is familiar with from the L1.

The present study has also revealed that many of the identified errors appear to have been triggered off by the presence of some other component which the learners tend to associate particularly strongly with a different tense than the one which they ought to

use. These errors may be teaching induced as in the grammar presentation and practice components of language courses and textbooks such features are highlighted and often made the focus of intensive and repeated practice. Our findings here are comparable to Granger's (1999), who identified similar triggers in her study of advanced English writing by native speakers of French. Götz (2015), for comparison, does not deal with such triggers, and it must be admitted that they might be one of several possible explanations for these errors.

The findings of our study are limited in many ways. First and foremost, the very process of identification of errors is problematic (as, for example, different norms may apply to the so-called spoken grammar). The work on the double-checking of our corpus is not completed yet and consequently we cannot provide any information regarding inter-annotator agreement. Secondly, despite the laboriousness involved in their compilation, spoken learner corpora are usually much smaller than written ones. This means that we cannot be quite certain whether the recorded instances are examples of systematic or random behaviour. Besides, LINDSEI subcorpora have not yet been morphologically tagged (although the work has commenced now), which makes it virtually impossible to perform full-scale analyses of the uses of individual grammatical features, for example comparing the erroneous and correct use of tenses. Thus our study focuses only on errors and not on the correct usage. This is an area which must be addressed in future research. Last but not least, the production is affected by the design of the task: thus owing to its design, LINDSEI may not have provided the speakers with as much opportunity to use the future tense (or refer to the future using other grammatical constructions) whilst the opportunities to use the past or present perfect tenses are plentiful. Despite all this uncertainty, we find that our analyses are in agreement with our teaching experience and results published in similar studies elsewhere. They are also some of the most recent analyses of advanced English spoken by Czech learners and may thus provide an important point of comparison for future studies.

5. Conclusion

The analysis has shown that even at an advanced level there exist aspects of grammar which are particularly error-prone. These are especially the use of tenses (especially the present perfect), tense agreements and conditionals. Whilst in only a small number of instances the errors here led to reduced intelligibility or even altered meaning, advanced learners – who were in this case students of English philology and thus aspiring to become experts in the field – would no doubt like to think of their English speech performances as not being marred by such errors.

The majority of the errors could be explained either by the effect of negative L1 transfer or by the existence of various triggers or false associations. Our findings thus have pedagogical implications which point in two possible directions: we recommend that work with advanced learners should systematically target these error-prone areas not leaving the L1 aside, and we also call for a reconsideration of how these problematic areas are approached in the teaching process, whether in practice or in the explanation and presentation of grammar aspects.

We strongly advocate the following principles: 1) discourse-driven approaches to the presentation of grammar so that learners become aware of the importance and role of context upon the selection of grammatical features; 2) moving away from decontextualized sentence-based presentation and practice of new points of grammar so that learners become accustomed to choosing aspects of grammar depending on the prevailing notion (such as cause and effect) rather than by clinging to overt features (such as the various triggers); 3) data-driven learning so that learners may benefit from analysing typical performance problems of other students of the same level of proficiency and learn from each other's mistakes; and 4) using L1-informed approaches to presenting new language material as these might help – if they are well designed – to avoid transfer related problems.

Whilst our study is limited in the number of learners (n = 50) and the size of the corpus (c. 125,000 tokens), it analyzes the largest collection of advanced spoken English produced by native Czech speakers available to date. As such it might be of benefit to language teachers and also researchers who should not only extend this research to other areas of grammar and lexis but also make a full account of those features which these learners find unproblematic.

REFERENCES

Anthony, L. (2014) *AntConc (Version 3.4.3)*. Tokyo: Waseda University.

Carlsen, C. (2012) Proficiency level – a fuzzy variable in computer learner corpora. *Applied Linguistics* 33/2, 161–183.

Clark, E. V. (2009) *First Language Acquisition*. Cambridge: University Press.

Conrad, S. (2016) Using corpus linguistics to improve the teaching of grammar. In: Hinkel, E. (ed.), *Teaching English Grammar to Speakers of Other Languages*. New York: Routledge.

Corder, S. P. (1967) The significance of learners' errors. *International Review of Applied Linguistics* 5, 161–170.

Corder, S. P. (1974) Error analysis. In: Allen, J. P. B., and S. P. Corder (eds.), *Techniques in Applied Linguistics*. London/New York: Oxford University Press.

Council of Europe (2001) *Common European Framework of Reference for Languages. Learning, teaching, assessment*. Cambridge: University Press.

Dagneaux, E., S. Denness and S. Granger (1998) Computer-aided error analysis. *System: An International Journal of Educational Technology and Applied Linguistics* 26/2, 163–174.

Dagneaux, E., S. Denness, S. Granger, F. Meunier, J. Neff Van Aertselaer and J. Thewissen (2008) *Error Tagging Manual Version* 1.3. Unpublished manual. Université catholique de Louvain.

Davydova, J. (2011) *The Present Perfect in Non-native Englishes: a Corpus-based Study of Variation*. Berlin/Boston: De Gruyter Mouton.

Dietrich, R., W. Klein and C. Noyau (1995) *The Acquisition of Temporality in a Second Language*. Amsterdam/Philadelphia: J. Benjamins.

Dušková, L. (1969) On sources of errors in foreign language learning. *International Review of Applied Linguistics* 7, 11–36.

Eriksson, A. (2004) *Tense and Aspect in Learner Writing. Advanced Swedish Learners' Use of Tense and Aspect in English Argumentative Text*. Göteborg: University of Göteborg.

Eriksson, A. (2008) *Tense and Aspect in Advanced Swedish Learners' Written English*. Göteborg: Acta Universitatis Gothoburgensis.

Gilquin, G. and S. De Cock (2011) Errors and disfluencies in spoken corpora: Setting the scene. *International Journal of Corpus Linguistics* 16/2, 141–172.

Gilquin, G., S. De Cock and S. Granger (2010) *The Louvain International Database of Spoken English Interlanguage. Handbook and CD-ROM*. Louvain-la-Neuve: Presses universitaires de Louvain.
Götz, S. (2015) Tense and aspect errors in spoken learner language: Implications for language testing and assessment. In Marcus, C. (ed.), *Learner Corpora in Language Testing and Assessment*, 191–215. Amsterdam/Philadelphia: J. Benjamins.
Gráf, T. (2015a) *Accuracy and Fluency in the Speech of the Advanced Learner of English* (Ph.D. Thesis). Prague: Charles University.
Gráf, T. (2015b) *Korpus LINDSEI_CZ*. Prague: Charles University.
Granger, S. (1996) From CA to CIA and back: an integrated approach to computerized bilingual and learner corpora. In: Aijmer, K., B. Altenberg and M. Johansson (eds.), *Languages in Contrast. Text-based Cross-linguistic Studies*, 37–51. Lund: Lund University Press.
Granger, S. (1999) Use of tenses by advanced EFL learners: evidence from an error-tagged computer corpus. In: Hasselgård, H. and S. Oksefjell (eds.), *Out of Corpora. Studies in Honour of Stig Johansson*, 191–202. Amsterdam/Atlanta: Rodopi.
Granger, S. (2009) The contribution of learner corpora to second language acquisition and foreign language teaching: A critical evaluation. In: Aijmer, K. (ed.), *Corpora and Language Teaching*. Amsterdam/Philadelphia: John Benjamins.
Granger, S. (2013) The passive in learner English: Corpus insights and implications for pedagogical grammar. In: Ishikawa, S. (ed.), *Papers from LCSAW2013* 1, 5–16. Japan: Kobe University.
Granger, S., F. Meunier and G. Gilquin (eds.) (2015) *The Cambridge Handbook of Learner Corpus Research*. Cambridge: University Press.
Guo, X. (2005) Modal auxiliaries in phraseology: a contrastive study of learner English and native speaker English. In: Danielsson, P. and M. Wagenmakers (eds.), *Proceedings from the Corpus Linguistics Conference Series* 1, Birmingham. Available at: http://www.birmingham.ac.uk/research/activity/corpus/publications/conference-archives/2005-birmingham.aspx.
Hinkel, E. (2004) Tense, aspect and the passive voice in L1 and L2 academic texts. *Language Teaching Research* 8/1, 5–29.
Housen, A., F. Kuiken and I. Vedder (eds.) (2012) *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency in SLA*. Amsterdam/Philadelphia: J. Benjamins.
Ingram, D. (1989) *First Language Acquisition: Method, Description and Explanation*. Cambridge: University Press.
James, C. (1998) *Errors in Language Learning and Use: Exploring Error Analysis*. London/New York: Longman.
Kammerer, S. (2009) Error-tagging spoken features of (learner) language: The UCL Error Editor revised. In: *Papers presented at the 30th annual conference of the International Computer Archive of Modern and Medieval English (ICAME 30)*. Lancaster.
Lado, R. (1957) *Linguistics Across Cultures: Applied Linguistics for Language Teachers.* Ann Arbor: University of Michigan Press.
Merilainen, L. (2010) *Language Transfer in the Written English of Finnish Students*. Joensuu: University of Eastern Finland.
Rogatcheva, S. I. (2009) I've only found the answer a few days ago: aspect use in Bulgarian and German EFL writing. In: Prado-Alonso, C., L. Gomez-Garcia, I. Pastor-Gomez and D. Tizon-Couto (eds.), *New Trends and Methodologies in Applied English Language Research. Diachronic, Diatopic and Contrastive Studies* 103, 255–278. Berlin: Peter Lang.
Rogatcheva, S. I. (2012) Measuring learner (mis)use: Tense and aspect errors in the Bulgarian and German components of ICLE. In: Thomas, J. and A. Boulton (eds.), *Input, Process and Product: Developments in Teaching and Language Corpora*, 258–272. Brno: Masaryk University.
Rundell, M. (2007) *Macmillan English Dictionary for Advanced Learners*. Oxford: Macmillan.
Salaberry, M. R., Y. Shirai and American Association for Applied Linguistics (eds.) (2002) *The L2 Acquisition of Tense-aspect Morphology*. Amsterdam: J. Benjamins.
Swan, M. and B. Smith (eds.) (2001) *Learner English: A Teacher's Guide to Interference and Other Problems* (2nd ed.). Cambridge, New York: Cambridge University Press.
Turton, N. D. and J. B. Heaton (1996) *Longman Dictionary of Common Errors*. Harlow: Longman.

CHYBY V UŽÍVÁNÍ ČASŮ V MLUVENÉ ANGLIČTINĚ POKROČILÝCH MLUVČÍCH

Resumé

Cílem studie bylo prozkoumat mluvený jazyk pokročilých mluvčích angličtiny s ohledem na chyby a nepřesnosti, jichž se dopouštějí v užívání morfologických kategorií sloves. Zdrojem dat pro studii byl korpus pokročilé žákovské angličtiny LINDSEI_CZ, který obsahuje transkribované nahrávky 50 studentů vyšších ročníků anglistiky na FF UK. Metodou, který byla využita, byla počítačová analýza chyb (Computer-aided Error Analysis). Její výsledky byly zpracovány kvantitativně, aby mohly být rozpoznány ty gramatické jevy, které pro pokročilé studenty představují nejčastější problém. Ukazuje se, že i takto pokročilí studenti chybují v poměrně základních užitích časů a časové souslednosti, a to především v užívání předpřítomného času. Vedle toho studie identifikuje i řadu jiných typických problémů. Studie poukazuje na to, že tyto problémy mohou pramenit z negativního transferu z mateřštiny, ale mohou být rovněž důsledkem nevhodných technik využívaných při vyučování těchto jevů. Řada příkladů totiž obsahuje společné prvky, které získávají pozornost při výkladu a procvičování gramatiky na úkor pěstování hlubšího pochopení takových sémantických konceptů jako je průběhovost či „perfektivita" (perfectiveness). Rozhodování studentů při volbě správného gramatického jevu pak může být ovlivněno přítomností marginálního prvku (např. adverbia, které se hojně využívá v určitém jevu), jež pak slouží jako spouštěč chyby, spíše než vyhodnocením kontextu a situačního zasazení promluvy. Studie následně vyzývá k přehodnocení postupů užívaných při vyučování gramatiky, a to především k odklonu od dekontextualizovaného procvičování a výkladu gramatiky na základě striktních pravidel a k příklonu k diskurzně orientovanému procvičování a vytváření hlubšího pochopení složitých konceptů ovlivňujících volbu času a vidu.

Tomáš Gráf
Department of English Language and ELT Methodology
Faculty of Arts, Charles University
Tomas.Graf@ff.cuni.cz